

Incorporating Signals into Optimal Trading

Charles-Albert Lehalle ^{1,2} and Eyal Neuman ^{*2,3}

¹Capital Fund Management, Paris

²CFM-Imperial College Institute, London

³Department of Mathematics, Imperial College London

April 7, 2017

Abstract

Optimal trading is a recent field of research which was initiated by Almgren, Chriss, Bertsimas and Lo in the late 90's. Its main application is slicing large trading orders, in the interest of minimizing trading costs and potential perturbations of price dynamics due to liquidity shocks. The initial optimization frameworks were based on mean-variance minimization for the trading costs. In the past 15 years, finer modelling of price dynamics, more realistic control variables and different cost functionals were developed. The inclusion of signals (i.e. short term predictors of price dynamics) in optimal trading is a recent development and it is also the subject of this work.

We incorporate a Markovian signal in the optimal trading framework which was initially proposed by Gatheral, Schied, and Slynko [20] and provide results on the existence and uniqueness of an optimal trading strategy. Moreover, we derive an explicit singular optimal strategy for the special case of an Ornstein-Uhlenbeck signal and an exponentially decaying transient market impact. The combination of a mean-reverting signal along with a market impact decay is of special interest, since they affect the short term price variations in opposite directions.

Later, we show that in the asymptotic limit where the transient market impact becomes instantaneous, the optimal strategy becomes continuous. This result is compatible with the optimal trading framework which was proposed by Cartea and Jaimungal [10].

*<http://eyaln13.wixsite.com/eyal-neuman>

1 Introduction

The financial crisis of 2008-2009 raised concerns about the inventories kept by intermediaries. Regulators and policy makers took advantage of two main regulatory changes (Reg NMS in the US and MiFID in Europe) which were followed by the creation of worldwide trade repositories. They also enforced more transparency on the transactions and hence on market participants positions, which pushed the trading processes toward electronic platforms [26]. Simultaneously, consumers and producers of financial products asked for less complexity and more transparency.

This tremendous pressure on the business habits of the financial system, shifted it from a customized and high margins industry, in which intermediaries could keep large (and potentially risky) inventories, to a mass market industry where logistics have a central role. As a result, investment banks nowadays unwind their risks as fast as possible. In the context of small margins and high velocity of position changes, trading costs are of paramount importance. A major factor of the trading costs is the market impact: the faster the trading rate, the more the buying or selling pressure will move the price in a detrimental way.

Academic efforts to reduce the transaction costs of large trades started with the seminal papers of Almgren and Chriss [5] and Bertsimas and Lo [8]. Both models deal with the trading process of one large market participant (for instance an asset manager or a bank) who would like to buy or sell a large amount of shares or contracts during a specified duration. The cost minimization problem turned out to be quite involved, due to multiple constraints on the trading strategies. On one hand, the market impact (see [7] are references therein) demands to trade slowly, or at least at a pace which takes into account the available liquidity. On the other hand, traders have an incentive to trade rapidly, because they do not want to carry the risk of an adverse price move far away from their decision price.

The importance of optimal trading in the industry generated a lot of variations for the initial mean-variance minimization of the trading costs (see [26, 15, 21] for details). In this paper, we consider the mean-variance minimization problem in the context of stochastic control (see e.g. [25], [9]). In this approach some more realistic control variables which are related to order book dynamics and specific stochastic processes for the underlying price can be used (see [22] and [29] for related work).

In this paper we address the question of how to incorporate *signals*, which are predicting short term price moves, into optimal trading problems. Usually optimal execution problems focus on the tradeoff between market impact and market risk. However, in practice many traders and trading algorithms use short term price predictors. Most of such documented predictors relate to orderbook dynamics [28]. They can be divided into two categories: signals which are based on liquidity consuming flows [11], and signals that measure the imbalance of the current liquidity. In [27], an example of how to use liquidity imbalance signals within a very short trading tactic is

studied. These two types of signals are closely related, since within short terms, price moves are driven by matching of liquidity supply and demand (i.e. current offers and consuming flows).

As mentioned earlier, one of the major influencers on transaction costs is the market impact. Empirical studies have shown that the influence of the market impact is *transient*, that is, it decays within a short time period after each trade (see [7] and references therein). In this paper we will focus on two frameworks which take into account different types of market impact:

- Gatheral, Schied and Slynko (GSS) framework [20], in which the market impact is transient and strategies have a fuel constraint, i.e., orders are finished before a given date T ;
- Cartea and Jaimungal (CJ) framework [10], where the market impact is instantaneous and the fuel constraint on the strategies is replaced by a smooth terminal penalization.

Note that [20] is not the only framework with market impact decay. This kind of dynamics was originally introduced in [30] and reused in [2] as in some other papers.

The main novelty which is introduced by our work is the use of a Markovian signal in the optimal trading problem within the (GSS) framework. We formulate a cost functional which consists of the trading costs and the risk of holding inventory at each given time. Then we prove that there exists at most one optimal strategy that minimizes this cost functional. The optimal strategy is formulated as a solution to a stochastic integral equation. We then derive explicitly the optimal strategy, for the special case where the signal is an Ornstein-Uhlenbeck process.

The use of a signal in optimal trading is relatively new (see [11] and [6]). To the best of our knowledge, this is the first time that a Markovian signal and a transient market impact are confronted. The (GSS) framework already includes a transient market impact, without using signals. The (CJ) framework includes only a bounded Markovian signal and not a decaying market impact. Moreover, our results on optimal trading in the (GSS) framework incorporate a risk eversion term in the cost functional, which was not taken into account in the results of [20].

Later in this paper we use the (CJ) framework for a qualitative comparison. We show that in the asymptotic regime where the transient market impact becomes instantaneous, the optimal strategy becomes continuous. Moreover, the asymptotics of the optimal strategy coincide with the optimal strategy within (CJ) framework. The comparison between different trading frameworks provides researchers and practitioners a wider overview, when they are facing real trading processes.

In order to validate our assumptions and theoretical results, we use real data from nordic European equity markets (the NASDAQ OMX exchange) to demonstrate the existence of a liquidity driven signal. We also show that practitioners are at least

partly conditioning their trading rate on this signal. Up to 2014, this exchange provided with each transaction the identity of the buyer and the seller. This database was already used for some academic studies, hence the reader can refer to [34] for more details. We added to these labelled trades, a database of Capital Fund Management (CFM) that contains information on the state of the order book just before each transaction. Thanks to this hybrid database, we were able to compute the imbalance of the liquidity just before decisions are taken by participants (i.e. sending a market orders which consume liquidity).

We divide most members of the NASDAQ OMX into four classes: global investment banks, institutional brokers, high frequency market makers and high frequency proprietary traders (the classification is detailed in the Appendix). Then, we compute the average value of the imbalance just before each type of participant takes a decision (see Figure 2). The conclusion is that some participants condition their trading rate on the liquidity imbalance. Moreover, we provide a few graphs that demonstrate a positive correlation between the state of the imbalance and the future price move. These graphs also provide evidences for the mean-reverting nature of the imbalance signal (see Figure 1). In Figure 3 we preset the estimated trading speed of market participants as a function of the average value of the imbalance, within a medium time scale of 10 minutes. The exhibited relation between the trading rate and the signal in this graph is compatible with our theoretical findings.

This paper is structured as follows. In Section 2 we introduce a model with a market impact decay, a Markovian signal and strategies with a fuel constraint. We provide general existence and uniqueness theorems, and then give an explicit solution for the case of an Ornstein-Uhlenbeck signal. The addition of a signal to market impact decay is the central ingredient of this Section. In Section 3 we show that the optimal strategy in the (GSS) framework coincides with the optimal strategy in the (CJ) framework, in the asymptotic limit where the transient market impact become instantaneous and the signal is an Ornstein-Uhlenbeck process. In Section 4 we provide an empirical evidence for the predictability of the imbalance signal and its use by different types of market participants. The last two sections are dedicated to proofs of the main results.

2 Model Setup and Main Results

2.1 Model setup and definition of the cost functional

In this section we define a model which incorporates a Markovian signal into the (GSS) optimal trading framework. Definitions and results from [20] are used throughout this section.

We consider a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfying the usual conditions,

where \mathcal{F}_0 is trivial. Let $M = \{M_t\}_{t \geq 0}$ is a right-continuous martingale and $I = \{I_t\}_{t \geq 0}$ a homogeneous continuous Markov process satisfying,

$$\int_0^T E_\iota[|I_t|] dt < \infty, \quad \text{for all } \iota \in \mathbb{R}, T > 0. \quad (2.1)$$

Here E_ι represents expectation conditioned on $I_0 = \iota$. In our model I represents a signal that is observed by the trader.

We assume that the asset price process P , which is unaffected by trading transactions is given by,

$$dP_t = I_t dt + dM_t, \quad t \geq 0,$$

hence the signal interacts with the price through the drift term. This setting allows us to consider a large class of signals. The visible asset price, which is described later, also depends on the market impact which is created by trader's transactions.

Let $[0, T]$ be a finite time horizon and $x > 0$ be the initial inventory of the trader. Let X_t be the amount of inventory held by the trader at time t . We define the class of admissible strategies $\Xi(x)$, such that every $X = \{X_t\}_{t \geq 0}$ in $\Xi(x)$ satisfies:

- (i) $t \longrightarrow X_t$ is left-continuous and (\mathcal{F}_t) -adapted.
- (ii) $t \longrightarrow X_t$ has finite and \mathbb{P} -a.s. bounded total variation.
- (iii) $X_0 = x$ and $X_t = 0$, \mathbb{P} -a.s. for all $t > T$.

As in [20, 17, 16], we assume that the visible price $S = \{S_t\}_{t \geq 0}$ is affected by a transient market impact, and it is given by

$$S_t = P_t + \int_{\{s < t\}} G(t-s) dX_s, \quad t \geq 0, \quad (2.2)$$

where the *decay kernel* $G : (0, \infty) \rightarrow [0, \infty)$ is a measurable function such that the following limit exists

$$G(0) := \lim_{t \downarrow 0} G(t). \quad (2.3)$$

Next we derive the transaction costs which are associated with the execution of a strategy X_t .

Note that if X_t is continuous in t , then the trading costs that arise by an infinitesimal order dX_t are $S_t dX_t$. When X_t has a jump of size ΔX_t at t , the price moves from S_t to $S_{t+} = S_t + G(0)\Delta X_t$ and the resulted costs by the trade ΔX_t are given by (see Section 2 of [20])

$$\frac{G(0)}{2} (\Delta X_t)^2 + S_t \Delta X_t.$$

It follows that the trading costs which arise from the strategy X are given by

$$\begin{aligned} \int S_t dX_t + \frac{G(0)}{2} \sum (\Delta X_t)^2 &= \int \int_0^t I_s ds dX_t + \int \int_{\{s < t\}} G(t-s) dX_s dX_t \\ &\quad + \int M_t dX_t + \frac{G(0)}{2} \sum (\Delta X_t)^2. \end{aligned}$$

From Lemma 2.3 in [20], we get a more convenient expression for the expected trading costs,

$$\begin{aligned} E \left[\int \int_0^t I_s ds dX_t + \int \int_{\{s < t\}} G(t-s) dX_s dX_t + \int M_t dX_t + \frac{G(0)}{2} \sum (\Delta X_t)^2 \right] \\ = E \left[\int \int_0^t I_s ds dX_t + \frac{1}{2} \int \int G(|t-s|) dX_s dX_t \right] - P_0 x. \end{aligned}$$

We are interested in adding a risk aversion term to the our cost functional. A natural candidate is $\int_0^T X_t^2 dt$, which is considered as a measure for the risk associated with holding the position X_t at time t ; see [4, 19, 33] and the discussion in Section 1.2 of [32]. Hence our cost functional which is the sum of the expected trading costs and the risk aversion term has the form

$$E \left[\int \int_0^t I_s ds dX_t + \frac{1}{2} \int \int G(|t-s|) dX_s dX_t + \phi \int_0^T X_t^2 dt \right] - P_0 x, \quad (2.4)$$

where $\phi \geq 0$ is a constant.

The main goal of this work is do minimize this cost functional (2.4) over the class of admissible strategies $\Xi(x)$. Before we discuss our main results in this framework, we introduce the following class of kernels.

We say that a continuous and bounded G is strictly positive definite if for every measurable strategy X we have

$$\int \int G(|t-s|) dX_s dX_t > 0, \quad P - \text{a.s.} \quad (2.5)$$

We define \mathbb{G} to be the class of continuous, bounded and strictly positive definite functions $G : (0, \infty) \rightarrow [0, \infty)$.

Remark 2.1. Note that (2.3) is satisfied for every $G \in \mathbb{G}$. A characterization of positive definite kernels (that is, when the inequality (2.5) is not strict) is given in Proposition 2.6 in [20].

Remark 2.2. An important subclass of \mathbb{G} is the class of bounded, non increasing convex functions $G : (0, \infty) \rightarrow [0, \infty)$ (see Proposition 2 in [3]).

2.2 Results for a Markovian Signal

In this section we introduce our results on the existence and uniqueness of an optimal strategy, when the signal is a continuous Markov process. In our first main result we prove that there exists at most one strategy which minimizes the cost functional (2.4).

Theorem 2.3. *Assume that $G \in \mathbb{G}$. Then, there exists at most one minimizer to the cost functional (2.4) in the class of admissible strategies $\Xi(x)$.*

In our next result we give a necessary and sufficient condition for the minimizer of cost functional.

Theorem 2.4. *$X^* \in \Xi(x)$ minimizes the cost functional (2.4) over $\Xi(x)$, if and only if there exists a constant λ such that X^* solves*

$$E \left[\int_0^t I_s ds + \int G(|t-s|) dX_s^* - 2\phi \int_0^t X_s^* ds \right] = \lambda, \quad \text{for all } 0 \leq t \leq T. \quad (2.6)$$

A few remarks are in order.

Remark 2.5. *In the spacial case where the agent does not rely on a signal (i.e. $I = 0$) and there is a zero risk aversion ($\phi = 0$), Theorems (2.3) and (2.4) coincide with Proposition 2.9 and Theorem 2.11 in [20].*

Remark 2.6. *Dang studied the case where the risk aversion term in (2.4) is nonzero, but again $I = 0$. In Section 4.2 of [17], a necessary condition for the existence of an optimal strategy is given, when the admissible strategies are deterministic and absolutely continuous. Our condition in (2.6) coincides with Dang's result when $I = 0$ and the admissible strategies are deterministic and absolutely continuous. Note however that the question if the condition in [17] is also sufficient and the uniqueness of the optimal solution, remained open even in the spacial case where $I = 0$.*

2.3 Result for an Ornstein-Uhlenbeck Signal

As mentioned in the introduction, a special attention is given to the case where the signal I is an Ornstein–Uhlenbeck process,

$$\begin{aligned} dI_t &= -\gamma I_t dt + \sigma dW_t, \quad t \geq 0, \\ I_0 &= \iota, \end{aligned} \quad (2.7)$$

where W is a standard Brownian motion and $\gamma, \sigma > 0$ are constants. In the following corollary we derive an explicit formula for the optimal strategy in the case of zero risk aversion and when G has an exponential decay. The following corollary generalizes the result of Obizhaeva and Wang [30], who solved this control problem when there no signal.

Corollary 2.7. *Let I be defined as in (2.7). Assume that $\phi = 0$ and $G(t) = \kappa \rho e^{-\rho t}$, where $\kappa, \rho > 0$ are constants. Then, there exists a unique minimizer $X^* \in \Xi(x)$ to the cost functional (2.4), which is given by*

$$X_t^* = x + \mathbf{1}_{\{t>0\}}A + \frac{B}{\gamma}(1 - e^{-\gamma t}) + Ct + \mathbf{1}_{\{t>T\}}D. \quad (2.8)$$

where

$$\begin{aligned} A &= \frac{1}{2 + T\rho} \left(\frac{\iota}{2\kappa\rho^2\gamma} \left((\rho + \gamma)(1 + T\rho + \gamma^{-1}(\rho - \gamma)(1 - e^{\gamma T})) - (\rho - \gamma)e^{-\gamma T} \right) - x \right), \\ B &= \iota \frac{\rho^2 - \gamma^2}{2\kappa\rho^2\gamma}, \\ C &= \rho A - \iota \frac{\rho + \gamma}{2\kappa\rho\gamma}, \\ D &= A - \frac{\iota}{2\kappa\rho^2\gamma} ((\rho + \gamma) - (\rho - \gamma)e^{-\gamma T}). \end{aligned}$$

Note that A, C, D are functions of (x, ι, T) while B is a function of ι .

Remark 2.8. *The optimal strategy X^* in this example is deterministic. We expect that in other cases the optimal strategies would not be typically deterministic, but would be adaptive in the sense of a nontrivial dependence in the signal and asset price (see for example the results in Section 3).*

Remark 2.9. *Note that in the limit where $\rho \rightarrow \infty$, the market impact term in (2.4), $\frac{1}{2} \int_0^T \int_0^T G(|t-s|) dX_s dX_t$ formally corresponds to the costs arising from instantaneous market impact, that is $G(dt) = \kappa \delta_0$. We briefly discuss the asymptotics of the optimal strategy $X_t^* = X_t^*(\rho)$ in (2.8) when $\rho \rightarrow \infty$. It is easy to verify that in the limit, the jumps of X^* , which are given by A and D , vanish and the limiting optimal strategy $X^*(\infty)$ is a smooth function which is given by*

$$X_t^*(\infty) = X + \frac{\iota}{2\kappa\gamma^2}(1 - e^{-\gamma t}) - \frac{\iota}{2\kappa\gamma}t.$$

Motivated by these asymptotic results, in the next section we further explore absolutely continuous strategies which minimize the trading costs – risk aversion functional. We will assume there that the market impact is instantaneous, that is $G(dt) = \kappa \delta_0$ and drop the fuel constraint ($X_t = 0$ for $t > T$) from the admissible strategies. Then, explicit formulas for the optimal strategy are derived when the risk aversion term is non-zero.

Organisation of the paper. In Section 3 we derive a closed form solutions for the case where the market impact is temporary (see Propositions (3.1) and (3.2)). In Section 4 we give an empirical evidence for trading which is based on a signal. The proofs of Theorems 2.3 and 2.4 and Corollary 2.7 are given in Section 5. In Section 6 we prove Propositions 3.1 and 3.2.

3 Optimal strategy for temporary market impact

In this section we study in greater detail the special case when the market impact is temporary, (i.e. $G(dt) = \kappa \delta_0(dt)$ for some $\kappa > 0$). We continue to assume that I is a continuous Markov process as in the beginning of Section 2 but we add the assumption that

$$E_\iota[|I_t|] \leq C(T)(1 + |\iota|), \quad \text{for all } \iota \in \mathbb{R}, \ 0 \leq t \leq T, \quad (3.1)$$

for some constant $C(T) > 0$.

For the sake of simplicity we will assume that $M_t = \sigma^P W_t$ so that

$$dP_t = I_t dt + \sigma^P dW_t,$$

where $\{W_t\}_{t \geq 0}$ is a Brownian motion and σ^P is a positive constant.

In the following example the fuel constraint on the admissible strategies will be replaced with a terminal penalty function. This allows us to consider absolutely continuous strategies as in the framework of Cartea and Jaimungal (see e.g. [12, 13, 14]). We introduce some additional definitions and notation which are relevant to this setting.

Let \mathcal{V} denote the class of progressively measurable control processes $r = \{r_t\}_{t \geq 0}$ for which $\int_0^T |r_t| dt < \infty$, P -a.s.

For any $x \geq 0$ we define

$$X_t^r = x - \int_0^t r_t dt. \quad (3.2)$$

Here X_t^r is the amount of inventory held by the trader at time t . We will often suppress the dependence of X in r , to ease the notation.

The price process, which is affected by the linear instantaneous market impact, is given by

$$S_t = P_t - \kappa r_t, \quad t \geq 0,$$

where $\kappa > 0$. Note that S_t here corresponds to (2.2) when $G(dt) = \kappa \delta_0(dt)$.

The investor's cash \mathcal{C}_t satisfies

$$d\mathcal{C}_t := S_t r_t dt = (P_t - \kappa r_t) r_t dt,$$

with $\mathcal{C}_0 = c$.

For the sake of consistency with earlier work of Cartea and Jaimungal in [12, 13, 14], we will define the liquidation problem as a maximization of the difference between the cash and the risk aversion. Moreover, the fuel constraint on the admissible strategies will be replaced by the penalty function $X_T(P_T - \varrho X_T)$ where ϱ is a positive constant.

The resulted cost functional is given by

$$V^r(t, \iota, c, x, p) = E_{\iota, c, x, p} \left[\mathcal{C}_T - \phi \int_t^T X_s^2 ds + X_T(P_T - \varrho X_T) \right], \quad (3.3)$$

where $\phi \geq 0$ is a constant and $E_{\iota, c, x, p}$ represents expectation conditioned on $I_t = \iota, \mathcal{C}_t = c, X_t = x, P_t = p$.

The value function is

$$V(t, \iota, c, x, p) = \sup_{r \in \mathcal{V}} V^r(t, \iota, c, x, p).$$

Note that this control problem could be easily transformed to a minimization of the trading costs and risk aversion as in Section 2.

Let \mathcal{L}^I be the generator of the process I . Then, the corresponding HJB equation is

$$0 = \partial_t V + \iota \partial_p V + \frac{1}{2}(\sigma^P)^2 \partial_p^2 V + \mathcal{L}^I V - \phi x^2 + \sup_r \left\{ r(p - \kappa r) \partial_c V - r \partial_x V \right\}, \quad (3.4)$$

with the terminal condition

$$V(T, \iota, c, x, p) = c + x(p - \varrho x).$$

In the following proposition we derive a solution to (3.4). The proof of Proposition 3.1 follows the same lines as the proof of Proposition 1 in [14].

Proposition 3.1. *Assume that $\varrho \neq \sqrt{\kappa\phi}$. Then, there exists a solution to (3.4) which is given by*

$$V(t, \iota, c, x, p) = c - xp + v_0(t, \iota) + xv_1(t, \iota) + x^2 v_2(t), \quad (3.5)$$

where

$$\begin{aligned} v_2(t) &= -\sqrt{\kappa\phi} \frac{1 + \zeta e^{2\beta(T-t)}}{1 - \zeta e^{2\beta(T-t)}}, \\ v_1(t, \iota) &= \int_t^T e^{\frac{1}{\kappa} \int_t^s v_2(u) du} E_{t, \iota}[I_s] ds, \\ v_0(t, \iota) &= \frac{1}{4\kappa} \int_t^T E_{t, \iota}[v_1^2(s, I_s)] ds, \end{aligned}$$

and the constants ζ and β are given by

$$\zeta = \frac{\varrho + \sqrt{\kappa\phi}}{\varrho - \sqrt{\kappa\phi}}, \quad \beta = \sqrt{\frac{\kappa}{\phi}}.$$

In the following proposition we prove that the solution to (3.4) is indeed an optimal control to (3.3).

Proposition 3.2. Assume that $\varrho \neq \sqrt{\kappa\phi}$. Then (3.5) maximizes the cost functional in (3.3). The optimal trading speed r_t^* is given by

$$r_t^* = -\frac{1}{2\kappa} \left(2v_2(t)X_t + \int_t^T e^{\frac{1}{\kappa} \int_t^s v_2(u)du} E_{t,t}[I_s] ds \right), \quad 0 \leq t \leq T.$$

The proofs of Propositions 3.1 and 3.2 are given in Section 6.

The following corollary follows directly from Propositions 3.1 and 3.2.

Corollary 3.3. Assume the same hypothesis as in Proposition 3.1, only now let I follow an Ornstein-Uhlenbeck process as in (2.7). Then, there exists a maximizer $r^* \in \mathcal{V}$ to $V^r(t, \iota, c, x, p)$, which is given by

$$r_t^* = -\frac{1}{2\kappa} \left(2v_2(t)X_t + I_t \int_t^T e^{-\gamma(s-t) + \frac{1}{\kappa} \int_t^s v_2(u)du} ds \right), \quad 0 \leq t \leq T.$$

In the following remarks we compare between the results of Sections 2 and 3.

Remark 3.4. If we set the risk aversion and penalty coefficients ϕ, ϱ in (3.3) to 0, then from the proof of Proposition 3.1, it follows that $v_2 \equiv 0$. Under the same assumptions on the signal as in Corollary 3.3, the optimal strategy is given by

$$r_t^* = -\frac{I_t}{2\kappa\gamma} (1 - e^{-\gamma(T-t)}), \quad 0 \leq t \leq T, \quad (3.6)$$

which is consistent with $X_t^*(\infty)$ from (2.9).

Remark 3.5. One can heuristically impose a “fuel constraint” on the optimal strategy in Corollary 3.3, by using the asymptotics of r_t^* when $\varrho \rightarrow \infty$. In this case $\zeta \rightarrow 1$ and the limiting optimal speed which we denote by r_t^f is

$$r_t^f = -\frac{1}{2\kappa} \left(2\bar{v}_2(t)X_t + I_t \int_t^T e^{-\gamma(s-t) + \frac{1}{\kappa} \int_t^s \bar{v}_2(u)du} ds \right), \quad 0 \leq t \leq T.$$

where

$$\bar{v}_2(t) = -\sqrt{\kappa\phi} \frac{1 + e^{2\beta(T-t)}}{1 - e^{2\beta(T-t)}}.$$

Remark 3.6. It is important to notice that (2.8) gives the optimal strategy on the time horizon $[0, T]$ in the (GSS) framework, by using only information on the O.U. signal at $t = 0$. On the other hand (3.6), which is the optimal trading speed r_t in the (CJ) framework is using the information on the signal at time t . A crucial point here is, that if one tries to solve repeatedly the control problem in the (GSS) framework on time intervals $[t, T]$ for any $t > 0$, by using I_t and S_t as an input, the optimal strategy will not necessarily minimize the cost functional (2.4) on $[0, T]$. The reason for that is, that the control problem in (2.4) is inconsistent because of the kernel $G(\cdot)$. The market impact (and therefore the transaction costs) which are created on $[0, t]$ effect the cost functional at $[t, T]$. This affect disappears when we set $G(dt) = \delta_t$ in (3.3).

4 Evidence for the use of signals in trading

In this section we provide an empirical evidence for the existence of a liquidity driven signal. We also study its dynamics and use by market participants. Note in Sections 2 and 3 we discussed more general signals which are not necessarily liquidity driven.

4.1 The database: NASDAQ OMX trades

The database which is used in this section is made of transactions on NASDAQ OMX exchange. This exchange used to publish the identity of the buyer and seller of each transaction until 2014. To obtain order book data, we use recordings made by Capital Fund Management (CFM) on the same exchange, which were matched with NASDAQ OMX trades thanks to the timestamp, quantity and price of each trade. On a typical month, the accuracy of such matching is more than 99.95%.

The NASDAQ OMX trades were already used for academic studies (see [34] and [27] for details). Here we focus on data of a very liquid stock, AstraZeneca, from January 2013 to September 2013. The purpose of this section is not to conduct an extensive econometric study on this database; such work deserves a paper of its own. Our goal here is to show qualitative evidences for the existence of the order book imbalance signal and to study how market participants decisions depend on its value.

The NASDAQ OMX database contains the identity of the buyer and the seller *from the viewpoint of the exchange*, that is, the members of the exchange who made the transactions. Asset managers for example, are not direct members of the exchange. On the other hand, brokers, banks and some other specific market participants are members. We classify the market members into four types (for more details see Appendix A):

- Global investment banks (GIB);
- Institutional brokers (IB);
- High frequency market makers (HFMM);
- High frequency proprietary traders (HFPT).

We expect institutional brokers to execute orders for clients without taking additional risks (i.e. act as “pure agency brokers”). Such brokers often have medium size clients and local asset managers. They do not spend a lot of resources such as technology or quantitative analysts to study the microstructure and react fast to microscopic events.

Global Investment Banks can take risks at least on a fraction of their order flow. Most of them already had proprietary trading desks and high frequency trading activities in 2013 (i.e. during the recording of the data). They usually have large

international clients and have the capability to react to changes in the state of the order-book.

High frequency market makers are providing liquidity on both sides of the order book. They have a very good knowledge on market microstructure. As market makers, we expect them to focus on adverse selection, and not to keep large inventories. On the other hand, high frequency proprietary traders take their own risks in order to earn money, while taking profit of their knowledge of the order book dynamics.

Participant Class	Trade Type	Average imbalance	Number	Pct. Order Type
Global Banks	Limit	-0.37	62,838	50.83
	Market	0.54	60,795	49.17
HF MM	Limit	-0.35	23,555	89.65
	Market	0.53	2,720	10.35
HF Prop.	Limit	-0.30	23,203	40.50
	Market	0.49	34,081	59.50
Instit. Brokers	Limit	-0.58	5,436	29.42
	Market	0.41	13,042	70.58
Total			362,728	

Table 1: Descriptive statistics of market participants from January 2013 to September 2013 on AstraZeneca.

Table 1 shows descriptive statistics on the four types of participants. First note that the number of attributed trades amounts to 63% of the total number of trades. This is because we only classified traders that could be identified without a doubt, thanks to their web site, news and description on Bloomberg terminals. The data in Table 1 is compatible with our prior knowledge on the different classes of traders:

- HFMM trade far more with limit orders (90%), than with market orders;
- HFPT and IB place more market orders than limit orders;
- GIB have balanced order flows.

The average imbalance column in Table 1 is addressed in the next section.

4.2 The Imbalance Signal

The order book imbalance has been identified as one of the main drivers of liquidity dynamics. It plays an important role in order-book models and more specifically it drives the rate of insertions and cancellations of limit orders near the mid price (see [1, 24]). As an illustration of the theoretical results of this paper, we document here

the *imbalance signal* and its use by different types of participants. This signal is computed by using the quantity of the best bid Q_B and the best ask Q_A of the order book,

$$\text{Imb}(\tau) = \frac{Q_B(\tau) - Q_A(\tau)}{Q_B(\tau) + Q_A(\tau)},$$

just before the occurrence of a transaction at time τ^+ . Note AstraZeneca is a “*medium tick stock*” in the sense its bid-ask spread is on average 1.27 times the tick. This means that the liquidity at the best bid and ask gives a substantial information on the price pressure (see [23] for details about the role of the tick size in liquidity formation). For smaller tick stocks, several price levels need to be aggregated in order to obtain the same level of prediction for future price moves.

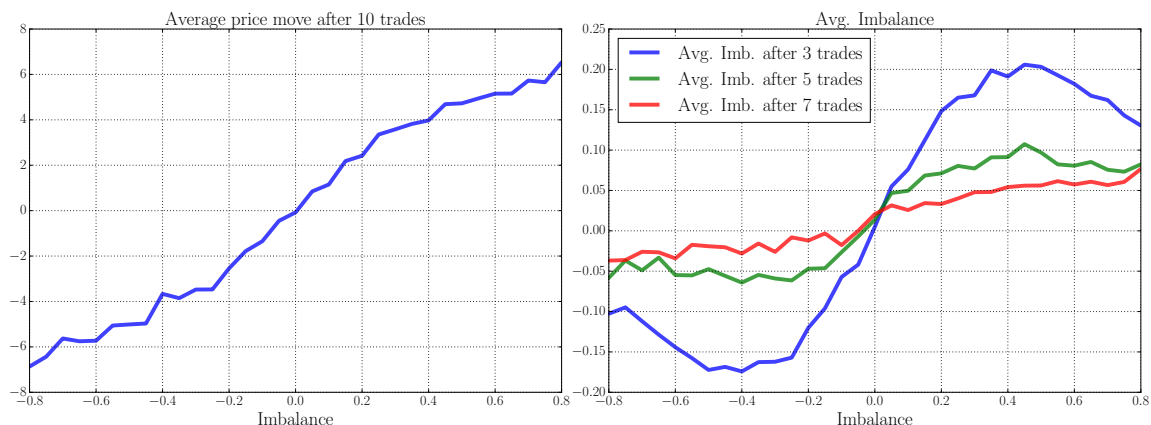


Figure 1: (Left) Predictive power of the imbalance: the average price move in the next 10 trades on the y -axis, as a function of the current imbalance on the x -axis. (Right) Mean-reversion of the imbalance: the average value of the imbalance after 3, 5 and 7 trades on the y -axis, as a function of the current imbalance on the x -axis.

In order to demonstrate the predictive power of the imbalance, we consider the average mid price move after 10 trades as a function of the current imbalance (see Figure 1 on the left). Keep in mind that we are not performing a detailed study of the imbalance signal. The charts and tables in this section are mainly informative and intend to justify our theoretical models.

Mean-reversion of the imbalance. Figure 1 on the right shows the average value of the imbalance after $\Delta T = 3, 5$ and 7 trades as a function of its current value. The decreasing slopes at $\text{Imb}(t) = 0$ as ΔT grows, demonstrate the mean reverting property of the imbalance. We will not comment too much on the decreasing slopes for large imbalance values. We will just mention that strong imbalance may imply on a future price change, which in turn, can create a depletion of the “weak side of the order-book” (in the sense of [18]). This phenomenon may cause an inversion of

the imbalance, since the queue in second best price level of the order book, which is now “promoted” to be the first level, could be large. See [24] for details about queue dynamics in order-books.

From Figure 1 we conclude the following:

- there is at least one liquidity driven short term signal, i.e. $\text{Imb}(\tau)$;
- this signal has mean-reverting properties;
- some market participants adjust their short term behaviour according to the value of $\text{Imb}(\tau)$.

4.3 Use of signals by market participants

As previously mentioned, we expect HF Proprietary Traders, HF Market Makers and Global Investment Banks to pay more attention to order-book dynamics than Institutional Brokers. However, as market makers, HFMM are expected to earn money by buying and selling when the mid price does not change much (relying on the *bid-ask bounce*). On the other hand, HFPT are typically alternating between intensive buy and sell phases which are based on price moves.

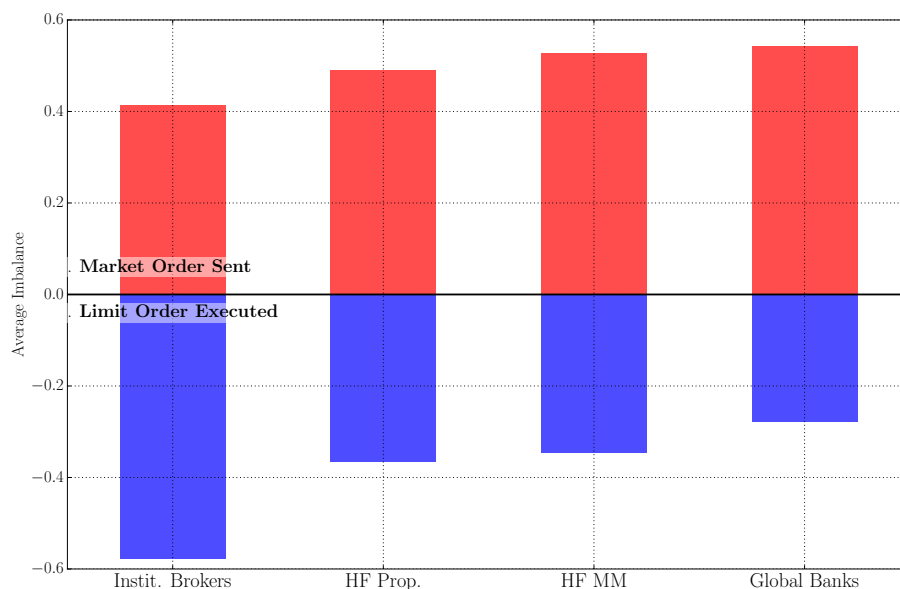


Figure 2: The average imbalance before a trade, using a limit buy order (blue) or a market buy order (red).

Our expectations are met in Figure 2, where the average imbalance just before a trade is shown for each type of market participants. All the bars are normalized as if

all orders were buy orders. The imbalance is positive when its sign is in the direction of the trade (positive for a buy order, or negative for a sell order). We notice the following behaviour:

- when the transaction is obtained via a market order (red bar), the market participant had the opportunity to observe the imbalance before consuming liquidity.
- when the transaction is obtained via a limit order (blue bar), fast participants have the opportunity to cancel their orders to prevent an execution and potential adverse selection.

Figure 2 underlines that HF participants and GBI make “better choices” on trading according to the market imbalance. Institutional Brokers seems to be the less “imbalance aware” when they decide to trade. This could be explained either by the fact that they invest less in microstructure research, quantitative modelling and automated trading; or either because they have less freedom to be opportunistic. Since they act as pure agency brokers, they do not have the choice to retain clients orders, and it could prevent them from waiting for the best imbalance to trade.

Strategic behaviour. Once we suspect that some participants take into account the imbalance in their trading decisions; we can look for a relation between the trading rate and the corresponding imbalance for each type of participant. This is motivated by the optimal trading frameworks of previous sections, where we used the trading rate as a control.

In order to learn more about the relation between the imbalance signal and the trading speed, we compute for all consecutive intervals of 10 minutes from January 2013 to September 2013 (during trading hours, i.e. 9h00 to 17h30):

- the average imbalance $\widehat{\text{Imb}}$ before a trade for each class of traders,
- the trading rates of the same participants when they buy \hat{r}_+ or sell \hat{r}_- . Here the trading rate is the number of shares bought (or sold) during 10 min, divided by the average number of shares bought or sold in 10 minutes over the whole database.

Figure 3 shows the connection between \hat{r}_\pm and the absolute value of the average imbalance $\widehat{\text{Imb}}$. The average imbalance appears on the x -axis. On the y -axis, the solid line represents the trading rate *in the direction of the imbalance* (i.e. buy order for a positive imbalance and sell orders for a negative one) and the dotted line is the trading rate *in the opposite direction of the imbalance*. The decreasing property of the curves comes from the fact that the number of high imbalance intervals is low. Therefore, any participant trades less during such intervals, than during intervals of

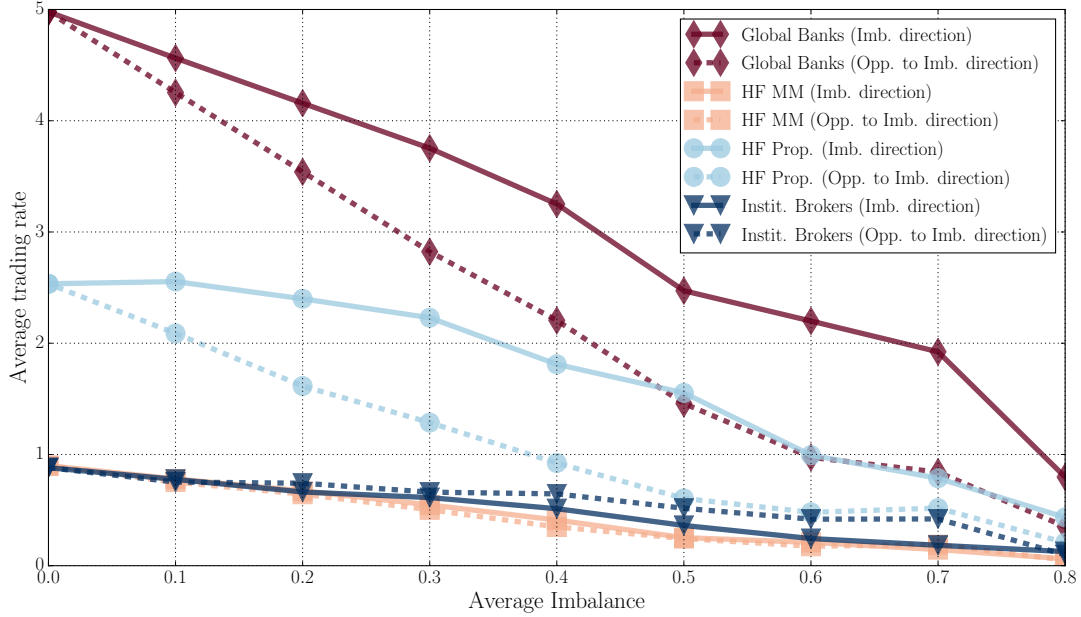


Figure 3: Average trading rate in the direction of the imbalance (solid line) and in the opposite direction (dotted line), during 10 consecutive minutes.

balanced signal which are more common. The important feature of Figure 3 is the difference between the dotted and the solid lines.

For a given average imbalance (i.e. one point on the x -axis), the solid line is the average trading speed of a given participant in the direction of the imbalance, while the dotted line is its average speed in the opposite direction. The difference between the dotted and the solid lines hence demonstrates the influence of the imbalance on the trading strategy for each class of participants. Figure 3 suggests that GIB and HFPT adjust their trading rates according to the imbalance, at the 10 minutes time scale. The HFMM, which seem to be imbalance aware at the time scale of their trades, do not exhibit a dependence at such long time scales. This could be explained by the fact that market makers cannot support a large inventory; they cannot sustain a significant difference between buys and sells during intervals as long as 10 minutes.

Towards a theory for the strategic use of signals. The analysis in this section suggests that some market participants are using liquidity-driven signals in their trading strategies. The liquidity imbalance, computed from the best bid and ask prices of the order-book for medium tick stocks, appears to be a good candidate. Moreover, its dynamics exhibit mean-reverting properties.

The theory developed in Sections 2 and 3 can be regarded as a tentative framework to model the behaviour the following participants. Global investment banks who

execute large orders, seem to be a typical example for participants who adopt the type of strategies that we model. HFPT who are combining slow signals (which may be considered as execution of large orders) along with fast signals, could also use our framework. We could moreover hope that thanks to the availability of such frameworks, Institutional Brokers could optimize their trading, and to increase the profits for more final investors.

5 Proofs of Theorems 2.3, 2.4 and Corollary 2.7

The proofs of Theorems 2.3 and 2.4 use ideas from the proofs of Proposition 2.9 and Theorem 2.11 in [20].

Proof of Theorem 2.3 Let $x \geq 0$. For any $X \in \Xi(x)$ define

$$C(X) := C_1(X) + C_2(X) + K(X), \quad (5.1)$$

where

$$\begin{aligned} C_1(X) &= \frac{1}{2} \int \int G(|t-s|) dX_s dX_t, \\ C_2(X) &= \phi \int_0^T X_s^2 ds, \\ K(X) &= \int \int_0^t I_s ds dX_t. \end{aligned}$$

Note that $E[(C(x))]$ is the cost functional in (2.4), where the constant $X_0 P_0$ is omitted.

Since $C_1(\cdot)$ is positive definite we have for any $X \in \Xi(x)$,

$$C_1(X) > 0. \quad (5.2)$$

$C_2(\cdot)$ is quadratic in X and therefore we have

$$C_2(X) \geq 0. \quad (5.3)$$

Let $X, Y \in \Xi(x)$. We define the following cross functionals,

$$\begin{aligned} C_1(X, Y) &= \frac{1}{2} \int \int G(|t-s|) dX_s dY_t, \\ C_2(X, Y) &= \phi \int_0^T X_s Y_s ds. \end{aligned}$$

Note that

$$C_i(X, Y) = C_i(Y, X), \quad \text{for } i = 1, 2,$$

and

$$C_i(X - Y) = C_i(X) + C_i(Y) - 2C_i(X, Y), \quad \text{for } i = 1, 2. \quad (5.4)$$

From (5.2) it follows that $C_1(X - Y) > 0$ and together with (5.4) we get

$$\begin{aligned} C_1\left(\frac{1}{2}X + \frac{1}{2}Y\right) &= \frac{1}{4}C_1(X) + \frac{1}{4}C_1(Y) + \frac{1}{2}C_1(X, Y) \\ &< \frac{1}{2}C_1(X) + \frac{1}{2}C_1(Y). \end{aligned}$$

Repeating the same steps, using (5.3) instead of (5.2) we get

$$C_2\left(\frac{1}{2}X + \frac{1}{2}Y\right) \leq \frac{1}{2}C_2(X) + \frac{1}{2}C_2(Y).$$

Since $K(X)$ is linear in X we have

$$K\left(\frac{1}{2}X + \frac{1}{2}Y\right) = \frac{1}{2}K(X) + \frac{1}{2}K(Y).$$

From (5.1) it follows that

$$C\left(\frac{1}{2}X + \frac{1}{2}Y\right) < \frac{1}{2}C(X) + \frac{1}{2}C(Y).$$

Let $\alpha \in (0, 1)$. The claim that

$$C(\alpha X + (1 - \alpha)Y) < \alpha C(X) + (1 - \alpha)C(Y),$$

follows from the continuity of $C(\cdot)$, by a standard extension argument. Since $C(\cdot)$ and therefore $E[C(X)]$ are strictly convex, we get that there exists at most one minimizer to $E[C(X)]$ in $\Xi(x)$. □

Proof of Theorem 2.4 First we prove that (2.6) is necessary for optimality. Let $0 \leq t < t_0 \leq T$ and consider the round trip

$$dY_s = \delta_{t_0}(ds) - \delta_t(ds).$$

For all $\alpha \in \mathbb{R}$ we have

$$C_i(X^* + \alpha Y) = C_i(X^*) + \alpha^2 C_i(Y) + 2\alpha C_i(X, Y), \quad i = 1, 2, \quad (5.5)$$

and

$$K(X^* + \alpha Y) = K(X^*) + \alpha K(Y) \quad (5.6)$$

Let $Z := X^* + \alpha Y$, and recall that $C(Z) = C_1(Z) + C_2(Z) + K(Z)$. Using (5.5) and (5.6) we can differentiate $E[C(Z)]$ with respect to α and get

$$\frac{\partial E[C(Z)]}{\partial \alpha} = E[K(Y)] + \sum_{i=1,2} 2\alpha E[C_i(Y)] + 2E[C_i(X, Y)].$$

From optimality we have $E[C(X^*)] \leq E[C(Z)]$ and therefore we expect that

$$\left. \frac{\partial E[C(Z)]}{\partial \alpha} \right|_{\alpha=0} = E[K(Y)] + 2 \sum_{i=1,2} E[C_i(X, Y)] = 0. \quad (5.7)$$

Note that

$$\begin{aligned} C_1(X, Y) &= \frac{1}{2} \int \int G(|t - s|) dX_s dY_t \\ &= \frac{1}{2} \int G(|t_0 - s|) dX_t^* - \frac{1}{2} \int G(|t - s|) dX_t^*, \end{aligned}$$

$$\begin{aligned} C_2(X, Y) &= \phi \int_0^T X_s Y_s ds \\ &= -\phi \int_t^{t_0} X_s ds. \end{aligned}$$

and

$$\begin{aligned} K(Y) &= \int \int_0^t I_s ds dY_t \\ &= \int_t^{t_0} I_s ds. \end{aligned}$$

We get that (5.7) is equivalent to

$$\begin{aligned} &E \left[\int G(|t_0 - s|) dX_t^* - 2\phi \int_0^{t_0} X_s ds + \int_0^{t_0} I_s ds \right] \\ &= E \left[\int G(|t - s|) dX_t^* - 2\phi \int_0^t X_s ds + \int_0^t I_s ds \right]. \end{aligned}$$

Since t and t_0 were chosen arbitrarily this implies (2.6).

Assume now that there exists $X^* \in \Xi(x)$ satisfying (2.6), we will show that X^* minimizes $E[C(\cdot)]$. Let X be any other strategy in $\Xi(x)$. Define $Z = X - X^*$. Then

from (2.6) we have

$$\begin{aligned}
E[C_1(X^*, Z)] &= E\left[\frac{1}{2} \int \int G(|t-s|) dX_s^* dZ_t\right] \\
&= E\left[\frac{1}{2} \int \left(\lambda + 2\phi \int_0^t X_s^* ds - \int_0^t I_s ds\right) dZ_t\right] \\
&= E\left[\frac{\lambda}{2} (X([0, \infty)) - X^*([0, \infty))) + \phi \int \int_0^t X_s^* ds dZ_t - \frac{1}{2} \int \int_0^t I_s ds dZ_t\right] \\
&= E\left[\phi \int \int_0^t X_s^* ds dZ_t - \frac{1}{2} K(Z)\right],
\end{aligned} \tag{5.8}$$

where we have used the fact that $X([0, \infty)) = X^*([0, \infty)) = x$ in the last equality.

From (5.5) and (5.8) we have

$$\begin{aligned}
E[C_1(X)] &= E[C_1(Z + X^*)] \\
&= E[C_1(Z) + C_1(X^*) + 2C_1(X^*, Z)] \\
&= E\left[C_1(Z) + C_1(X^*) - K(Z) + 2\phi \int \int_0^t X_s^* ds dZ_t\right],
\end{aligned}$$

and

$$\begin{aligned}
C_2(X) &= C_2(Z + X^*) \\
&= C_2(Z) + C_2(X^*) + 2C_2(X^*, Z) \\
&= C_2(Z) + C_2(X^*) + 2\phi \int_0^T X_s^* Z_s ds.
\end{aligned}$$

From the linearity of $K(\cdot)$ we get

$$K(X) = K(Z) + K(X^*).$$

It follows that

$$\begin{aligned}
E[C(X)] &= \sum_{i=1,2} E[C_i(X)] + E[K(X)] \\
&= E\left[C_1(X^*) + C_2(X^*) + K(X^*) + C_1(Z) + C_2(Z) \right. \\
&\quad \left. + 2\phi \int \int_0^t X_s^* ds dZ_t + 2\phi \int_0^T X_s^* Z_s ds\right] \\
&= E\left[C(X^*) + C_1(Z) + C_2(Z) \right. \\
&\quad \left. + 2\phi \int \int_0^t X_s^* ds dZ_t + 2\phi \int_0^T X_s^* Z_s ds\right].
\end{aligned}$$

Racal that $Z_0 = 0$ and $Z_t = 0$ for every $t > T$, hence from integration by parts we have

$$0 = \int_0^t \int_0^s X_s dZ_t + \int_0^T X_t Z_t dt,$$

and since for $i = 1, 2$, $E[C_i(Z)] \geq 0$, we get

$$E[C(X)] \geq E[C(X^*)].$$

□

Proof of Corollary 2.7 From (2.7) it follows that $E_t[I_t] = \iota e^{-\gamma t}$. Since $\phi = 0$, (2.6) reduces to

$$\frac{\iota}{\gamma}(1 - e^{-\gamma t}) + \kappa\rho \int_0^T e^{-\rho|t-s|} dX_s = \lambda. \quad (5.9)$$

Moreover we have the fuel constraint,

$$\int_0^T dX_t = -x. \quad (5.10)$$

Motivated by the example in Obizhaeva and Wang [30], we guess a solution of the form

$$dX_t = A\delta_0 + (Be^{-\gamma t} + C)dt + D\delta_T, \quad (5.11)$$

where δ_x is the Dirac's delta measure at x and A, B, C, D are some constants.

Note that

$$\kappa\rho \int_0^t e^{-\gamma s} e^{-\rho(t-s)} ds = \frac{\kappa\rho}{\rho - \gamma} (e^{-\gamma t} - e^{-\rho t}),$$

$$\kappa\rho \int_t^T e^{-\gamma s} e^{-\rho(s-t)} ds = \frac{\kappa\rho}{\rho + \gamma} (e^{-\gamma t} - e^{-\gamma T - \rho(T-t)}),$$

and therefore

$$\begin{aligned} & \kappa\rho \int_0^T e^{-\rho|t-s|} dX_s \\ &= \kappa\rho e^{-\rho t} A + B \frac{\kappa\rho}{\rho - \gamma} (e^{-\gamma t} - e^{-\rho t}) + B \frac{\kappa\rho}{\rho + \gamma} (e^{-\gamma t} - e^{-\gamma T - \rho(T-t)}) \\ & \quad + C\kappa(1 - e^{-\rho t}) + C\kappa(1 - e^{-\rho(T-t)}) + D\kappa\rho e^{-\rho(T-t)}. \end{aligned}$$

If we choose

$$\lambda = 2\kappa C + \frac{\iota}{\gamma},$$

we get from (5.9) and (5.10) the following linear system,

$$\begin{aligned} -\frac{\iota}{\gamma}e^{-\gamma t} + B\frac{\kappa\rho}{\rho-\gamma}e^{-\gamma t} + B\frac{\kappa\rho}{\rho+\gamma}e^{-\gamma t} &= 0, \\ A\kappa\rho e^{-\rho t} - B\frac{\kappa\rho}{\rho-\gamma}e^{-\rho t} - C\kappa e^{-\rho t} &= 0, \\ -B\frac{\kappa\rho}{\rho+\gamma}e^{-\gamma T-\rho(T-t)} - C\kappa e^{-\rho(T-t)} + D\kappa\rho e^{-\rho(T-t)} &= 0, \\ A + \frac{B}{\gamma}(1 - e^{-\gamma T}) + CT + D &= -x. \end{aligned}$$

From the first equation we can get

$$B = \iota \frac{\rho^2 - \gamma^2}{2\kappa\rho^2\gamma}.$$

and then

$$\begin{aligned} A &= \frac{1}{2+T\rho} \left(\frac{\iota}{2\kappa\rho^2\gamma} \left((\rho+\gamma)(1+T\rho+\gamma^{-1}(\rho-\gamma)(1-e^{\gamma T})) - (\rho-\gamma)e^{-\gamma T} \right) - x \right), \\ C &= \rho A - \iota \frac{\rho+\gamma}{2\kappa\rho\gamma}, \\ D &= A - \frac{\iota}{2\kappa\rho^2\gamma} ((\rho+\gamma) - (\rho-\gamma)e^{-\gamma T}). \end{aligned}$$

The optimal strategy is therefore

$$X_t^* = x + \mathbb{1}_{\{t>0\}}A + Ct + \frac{B}{\gamma}(1 - e^{-\gamma t}) + \mathbb{1}_{\{t>T\}}D.$$

□

6 Proofs of Propositions 3.1 and 3.2

Proof of Proposition 3.1. The proof follows the same lines as the proof of Proposition 1 in [14].

Pluggin in the ansatz $V(t, \iota, c, x, p) := c + xp + v(t, x, \iota)$ we get

$$0 = \partial_t v + \mathcal{L}^I v + \iota x - \phi x^2 + \sup_r \{ -r^2\kappa - r\partial_x v \}.$$

Optimizing over r it follows that

$$r^* = -\frac{\partial_x v}{2\kappa}, \quad (6.1)$$

and we get the following PDE

$$\partial_t v + \mathcal{L}^I v + \frac{1}{4\kappa} (\partial_x v)^2 + \iota x - \phi x^2 = 0. \quad (6.2)$$

where $v(T, x, \iota) = -\varrho x^2$.

As in equation (A.2) in [14], we have a linear and quadratic x -terms in (6.2) along with a quadratic terminal contention, hence we make the following anzats on the solution:

$$v(t, x, \iota) = v_0(t, \iota) + x v_1(t, \iota) + x^2 v_2(t, \iota).$$

By comparing terms with similar powers of q , we get the following system of PDEs,

$$\partial_t v_0 + \mathcal{L}^\iota v_0 + \frac{1}{4\kappa} v_1^2 = 0, \quad (6.3)$$

$$\partial_t v_1 + \mathcal{L}^\iota v_1 + \frac{1}{\kappa} v_2 v_1 + \iota = 0, \quad (6.4)$$

$$\partial_t v_2 + \mathcal{L}^\iota v_2 + \frac{1}{\kappa} v_2^2 - \phi = 0, \quad (6.5)$$

with the terminal conditions

$$v_0(T, \iota) = 0, \quad v_1(T, \iota) = 0, \quad v_2(T, \iota) = -\varrho.$$

We first find a solution to (6.5). Note that since the terminal condition is independent of ι we might be able to find a ι independent solution, that is $v_2(t) := v_2(t, \iota)$ which satisfies

$$\partial_t v_2 + \frac{1}{\kappa} v_2^2 - \phi = 0.$$

This is a Riccati equation which has the following solution (see the proof of Proposition 1 in [14]),

$$v_2(t) = -\sqrt{\kappa\phi} \frac{1 + \zeta e^{2\beta(T-t)}}{1 - \zeta e^{2\beta(T-t)}},$$

where

$$\zeta = \frac{\varrho + \sqrt{\kappa\phi}}{\varrho - \sqrt{\kappa\phi}}, \quad \beta = \sqrt{\frac{\phi}{\kappa}}.$$

Using v_2 , we can find a Feynman-Kac representation to the solution of (6.4),

$$\begin{aligned} v_1(t, \iota) &= E_{t, \iota} \left[\int_t^T e^{\frac{1}{\kappa} \int_t^s v_2(u) du} I_s ds \right] \\ &= \int_t^T e^{\frac{1}{\kappa} \int_t^s v_2(u) du} E_{t, \iota} [I_s] ds. \end{aligned}$$

Again by Feynman-Kac formula we derive a solution to (6.3),

$$\begin{aligned} v_0(t, \iota) &= E_{t, \iota} \left[\frac{1}{4\kappa} \int_t^T v_1^2(s, I_s) ds \right] \\ &= \frac{1}{4\kappa} \int_t^T E_{t, \iota} [v_1^2(s, I_s)] ds. \end{aligned}$$

Proof of Proposition 3.2 Note that V is a classical solution to 3.4. By standard arguments (see e.g. Theorem 3.5.2 in [31]), in order to prove that V in (3.5) is the value function of (3.3), it is enough to show that r^* is admissible and that

$$|V(t, \iota, c, x, p)| \leq C(1 + \iota^2 + c^2 + x^2 + p^2), \quad \text{for all } t \geq 0, \iota, c, x, p \in \mathbb{R}. \quad (6.6)$$

Clearly $\sup_{t \in [0, T]} |v_2(t)| < \infty$. From our conditions on I we have

$$E_\iota [|I_t|] \leq C(1 + |\iota|), \quad \text{for all } \iota \in \mathbb{R}, 0 \leq t \leq T,$$

then we will have

$$\begin{aligned} |xv_1(t, \iota)| &\leq C|x|(1 + |\iota|) \\ &\leq C(1 + \iota^2 + x^2), \quad \text{for all } t \geq 0, \iota, x \in \mathbb{R}, \end{aligned}$$

$$|v_0(t, \iota)| \leq C(1 + \iota^2), \quad \text{for all } t \geq 0, \iota \in \mathbb{R}.$$

and (6.6) follows. To prove that r^* is admissible it is enough to show that $\int_0^T |r_t^*| dt < \infty$. Since v_2 is bounded we notice that

$$\begin{aligned} |r_t^*| &\leq \frac{1}{2\kappa} \left(2|v_2(t)||X_t| + \int_t^T e^{\frac{1}{\kappa} \int_t^s |v_2(u)| du} E_{t, \iota} [|I_s|] ds \right) \\ &\leq C_1 |X_t| + C_2 T(1 + |\iota|) \\ &\leq (C_2 + C_1)(x + T(1 + |\iota|)) + C_1 \int_0^t |r_s| ds, \end{aligned}$$

where we used (3.2) in the last inequality. From Gronwall inequality we have

$$|r_t^*| \leq (C_2 + C_1)(x + T(1 + |\iota|))e^{C_1 T},$$

hence r^* is admissible. □

References

- [1] F. Abergel, M. Anane, A. Chakraborti, A. Jedidi, and I. M. Toke. *Limit Order Books (Physics of Society: Econophysics and Sociophysics)*. Cambridge University Press, 1 edition, May 2016.
- [2] A Alfonsi, A. Fruth, and A. Schied. Optimal execution strategies in limit order books with general shape functions. *Quant. Finance*, 10:143–157, 2010.
- [3] A Alfonsi, A Schied, and A. Slynko. Order book resilience, price manipulation, and the positive portfolio problem. *SIAM J. Financial Math.*, 3(1):511–533, 2012.
- [4] A. Almgren. Optimal trading with stochastic liquidity and volatility. *SIAM J. Financial Math.*, 3:163–181, 2012.
- [5] R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3(2):5–39, 2000.
- [6] R. Almgren and J. Lorenz. Adaptive arrival price. *Institutional Investor Journals: Algorithmic Trading III*, Spring 2007.
- [7] E. Bacry, A. Luga, M. Lasnier, and C. A. Lehalle. Market Impacts and the Life Cycle of Investors Orders. *Market Microstructure and Liquidity*, 1(2), December 2015.
- [8] D. Bertsimas and A. W. Lo. Optimal control of execution costs. *Journal of Financial Markets*, 1(1):1–50, 1998.
- [9] B. Bouchard, N. M. Dang, and C. A. Lehalle. Optimal control of trading algorithms: a general impulse control approach. *SIAM J. Financial Mathematics*, 2(1):404–438, 2011.
- [10] Á. Cartea and S. Jaimungal. Modelling asset prices for algorithmic and high-frequency trading. *Applied Mathematical Finance*, 20(6):512–547, 2013.
- [11] Á. Cartea and S. Jaimungal. Incorporating Order-Flow into Optimal Execution. *Social Science Research Network Working Paper Series*, January 2015.
- [12] Á. Cartea and S. Jaimungal. Optimal execution with limit and market orders. *Quantitative Finance*, 15(8):1279–1291, 2015.
- [13] Á. Cartea and S. Jaimungal. Risk metrics and fine tuning of high-frequency trading strategies. *Mathematical Finance*, 25(3):576–611, 2015.
- [14] Á. Cartea and S. Jaimungal. Incorporating order-flow into optimal execution. *Mathematics and Financial Economics*, 10(3):339–364, 2016.

- [15] Á. Cartea, S. Jaimungal, and J. Penalva. *Algorithmic and High-Frequency Trading (Mathematics, Finance and Risk)*. Cambridge University Press, 1 edition, October 2015.
- [16] G. Curato, J. Gatheral, and F. Lillo. Optimal execution with non-linear transient market impact. *Quantitative Finance*, 17(1):41–54, 2017.
- [17] N. M. Dang. Optimal execution with transient impact. *ssrn-id2183685*, 2014.
- [18] Pietro Fodra and Huy  n Pham. Semi Markov model for market microstructure, May 2013.
- [19] P. Forsyth, J. Kennedy, T. S. Tse, and H. Windclif. Optimal trade execution: a mean-quadratic-variation approach. *Journal of Economic Dynamics and Control*, 36:1971–1991, 2012.
- [20] J. Gatheral, A. Schied, and A. Slynko. Transient linear price impact and Fredholm integral equations. *Math. Finance*, 22:445–474, 2012.
- [21] O. Gu  ant. *The Financial Mathematics of Market Liquidity: From Optimal Execution to Market Making*. Chapman and Hall/CRC, April 2016.
- [22] O. Gu  ant, C. A. Lehalle, and J. Fernandez-Tapia. Optimal Portfolio Liquidation with Limit Orders. *SIAM Journal on Financial Mathematics*, 13(1):740–764, 2012.
- [23] W. Huang, C. A. Lehalle, and M. Rosenbaum. How to predict the consequences of a tick value change? Evidence from the Tokyo Stock Exchange pilot program, July 2015.
- [24] Weibing Huang, Charles-Albert Lehalle, and Mathieu Rosenbaum. Simulating and analyzing order book data: The queue-reactive model. *Journal of the American Statistical Association*, 10(509), December 2015.
- [25] I. Kharroubi and H. Pham. Optimal portfolio liquidation with execution cost and risk. *SIAM Journal on Financial Mathematics*, 1(1):897–931, June 2010.
- [26] C. A. Lehalle, S. Laruelle, R. Burgot, S. Pelin, and M. Lasnier. *Market Microstructure in Practice*. World Scientific publishing, 2013.
- [27] C. A. Lehalle and O. Mounjid. Limit Order Strategic Placement with Adverse Selection Risk and the Role of Latency, October 2016.
- [28] A. Lipton, U. Pesavento, and M. G. Sotiropoulos. Trade arrival dynamics and quote imbalance in a limit order book, December 2013.

- [29] E. Neuman and A. Schied. Optimal portfolio liquidation in target zone models and catalytic superprocesses, 2015.
- [30] A. A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets*, 16(1):1 – 32, 2013.
- [31] H. Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2009.
- [32] A. Schied. A control problem with fuel constraint and Dawson–Watanabe superprocesses. *Ann. Appl. Probab.*, 23(6):2472–2499, 2013.
- [33] S. T. Tse, P. A. Forsyth, J. S. Kennedy, and H. Windcliff. Comparison between the mean-variance optimal and the mean-quadratic-variation optimal trading strategies. *Appl. Math. Finance*, 20(5):415–449, 2013.
- [34] V. van Kervel and A. Menkveld. Do High-Frequency Traders Engage in Predatory Trading? Technical report, VU University Amsterdam., October 2014.

A Composition of market participants groups

High Frequency Traders

Name	NASDAQ-OMX member code(s)	Market Maker	Prop. Trader
All Options International B.V.	AOI	Yes	
Hardcastle Trading AG	HCT		
IMC Trading B.V	IMC, IMA		
KCG Europe Limited	KEM, GEL		
MMX Trading B.V	MMX		
Nyenburgh Holding B.V.	NYE	Yes	Yes
Optiver VOF	OPV		
Spire Europe Limited	SRE, SREA, SREB		
SSW-Trading GmbH	IAT		
WEBB Traders B.V	WEB		
Wolverine Trading UK Ltd	WLV		

Table 2: Composition of the group of HFT used for empirical examples, and the composition of our “high frequency market maker” and “high frequency proprietary traders” subgroups.

Global Investment Banks

Name	NASADQ-OMX member code(s)
Barclays Capital Securities Limited Plc	BRC
Citigroup Global Markets Limited	SAB
Commerzbank AG	CBK
Deutsche Bank AG	DBL
HSBC Bank plc	HBC
Merrill Lynch International	MLI
Nomura International plc	NIP

Table 3: Composition of the group of Global Investment Banks used for empirical examples.

Institutional Brokers

Name	NASADQ-OMX member code(s)
ABG Sundal Collier ASA	ABC
Citadel Securities (Europe) Limited	CDG
Erik Penser Bankaktiebolag	EPB
Jefferies International Limited	JEF
Neonet Securities AB	NEO
Remium Nordic AB	REM
Timber Hill Europe AG	TMB

Table 4: Composition of the group of Institutional Brokers used for empirical examples.